DOCUMENT RESUME

ED 288 391 FL 017 031

AUTHOR Gazdar, Gerald; Pullum, Geoffrey K.
TITLE A Note on Subcategorization, Constituent Order, and the Notion "Head."
PUB DATE Nov 86
NOTE 9p.; In: York Papers in Linguistics No. 12, 1986; see FL 017 027.
PUB TYPE Reports - Evaluative/Feasibility (142) -- Viewpoints (120)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Classification; Foreign Countries; Grammatical Acceptability; *Linguistic Theory; *Phrase Structure; Research Methodology; *Structural Analysis (Linguistics); *Syntax; *Transformational Generative Grammar

ABSTRACT

The authors of a previous paper on aspects of generalized phrase structure grammar respond to criticism of that paper and clarify elements in the discussion. The original paper addressed the problem of expressing relevant generalizations about the order of complements that a lexical item subcategorizes for. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

S. Harlow

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED288391

FL017031

A NOTE ON SUBCATEGORIZATION, CONSTITUENT ORDER, AND THE NOTION 'HEAD'*

Gerald Gazdar and Geoffrey K Pullum

(School of Social Sciences, University of Sussex and
Cowell College, University of California, Santa Cruz)

In this note, we offer a brief comment on a recent discussion by Richard Hudson of our 1981 paper 'Subcategorization, constituent order, and the notion "head"'. Hudson says that the paper "makes three rather separate points" (1983: 446-447; all subsequent quotations from these pages unless otherwise specified), and he attempts to summarize and discuss these three points. Unfortunately, he completely misrepresents all of them. This makes his discussion a strikingly unreliable source of information about what we proposed in our paper. Because the latter provides the basis for a significant amount of published work employing the 'ID/LP format' for grammars, it seems sensible to us to try and unravel Hudson's various misunderstandings.

1. Subcategorization

The first point Hudson sees in our paper is that we propose to

> give each construction relevant to subcategorization a number, so that instead of saying that a verb is transitive by saying it has the subcategorization feature [__NP], you call this construction (say) number 37, and then show that a verb is compatible with it by using [37] as a syntactic feature on the verb node.

He says it seems "extraordinary" to him "that this trick should have the appeal that it does have" for us; his reasoning runs as follows:

> If the numbers are just arbitrary diacritics, why not replace them by subcategorization features like [__NP], which are just copies of the constructions concerned - which takes you back to the standard view?

It would indeed be extraordinary if we imagined that there was an interesting difference between representing the features of intransitivity, transitivity, and ditransitivity as [37], [38], [39], rather than [__#], [__NP], [__NP PP] or [Intransitive], [Transitive], [Ditransitive]. The difference between what we propose and what "the

© the authors.

standard view" (that of Chomsky 1965, henceforth 'Aspects') assumes lies in the interpretation of the formalism in question. Features like [__NP PP], under the Aspects interpretation, are not atomic; they have internal structure. This internal structure is not merely a typographical nicety, a mnemonic for the working linguist: the internal structure plays an essential role in the Aspects theory of lexical insertion. Under this theory, a transformation examines the environment described by the internal structure of the feature and attempts to match it to a restricted class of environments in a tree.

By contrast, in the theory of subcategorization set out in our paper, a feature like [39] has no internal structure (e g it does not decompose into 3 and 9). It is not inspected by any lexical insertion transformation - indeed no such transformation exists. Nor does it, eo ipso, describe an environment in a tree. All it does is index a particular phrase structure rule. As Frank Heny has pointed out, in a passage that we quote in our paper, and as Hudson appears to accept, the Aspects theory is committed to massive redundancy, since "the lexicon will necessarily repeat information time and time again, which is at least in part already extractable from the PS rules" (Heny 1979: 339-340). This redundancy is completely avoided under our account since lexical entries simply contain pointers (to use a useful term from computer science) to phrase structure rules. The redundancy is not eliminated by any 'notational device'; rather it is eliminated by a radically different theory of how subcategorization works.

To see that Hudson is quite wrong in seeing this as a matter of notation, observe that there is a possibility inherent in our theory which does not exist in the Aspects theory. Under our theory, it is possible for there to be more than one rule in a grammar that corresponds to a single Aspects-style subcategorization frame. For example, we might want to have one rule 'VP --> H, VP' for verbs like 'try' and another, syntactically identical (until we consider the rule number), for verbs like 'tend', the two rules having different semantic translations.[1] Our theory can postulate two rules distinct only in their rule numbers in such a case; but the Aspects theory would represent the subcategorization property of a verb that took a VP complement by assigning it the feature [__VP]. There would be no way to differentiate 'try' from 'tend' without the introduction of new diacritic features (parallel to our rule number features, in fact) or other devices entirely outside the subcategorization mechanism. The same would be true for other such pairs. Note for example that 'see' and 'seem' both take a postverbal NP, but they do not have the same syntactic behaviour; 'see' passivizes (cf 'Jim sees everything'; 'Everything is seen by Jim') but 'seem' does not (cf 'Jim seems a nice guy'; *'A nice guy is seemed by Jim').

Although Hudson implies that our proposed change to the theory of subcategorization ("this trick") is merely notational - incorrectly, as we have just seen - he also claims that the Aspects notation has "descriptive advantages". Plainly, this is inconsistent. There can hardly be advantages inherent in the Aspects features but not in ours, if we have only departed from them notationally, as a "trick"; Hudson is adopting Chomsky's familiar rhetorical ploy of saying that on the one hand the proposal is a mere notational variant, and on the other hand that it is wrong. However, inconsistency aside, Hudson's argument for the superiority of the Aspects theory is also mistaken. Hudson argues thus: (i) The verb `joke' can be intransitive or can take one or two PPs as complements; (ii) we would represent this with a lexical entry such as <joke: 1, 3, 8>; (iii) the `standard theory' would represent the subcategorization possibilities of `joke' so as to express "the linguistically simple fact that `joke' takes up to two PP's" (p 447).

Points (i) and (ii) are correct, but (iii) cannot be accepted. It is notable that Hudson omits any `standard theory' representation of the entry for `joke' to illustrate his point. Such an entry will look like either <joke: [__ #], [__PP #], [__PP PP #]>, or <joke: [__(PP)(PP)#]>, depending on the conventions permitted (Chomsky himself handwaved on exactly this point: 1965: 111). To be sure, either one is more mnemonic than <joke: 1, 3, 8>; but that is a fact about the memory abilities of human beings (and human beings who are used to the formalism of Chomsky 1965, at that). Nothing linguistic follows from this mnemonicity in any way. For some reason, Hudson thinks that because the phrase `takes up to two PPs' occurs to him when he looks at the three Aspects features, this has something to do with the generalizations expressible in terms of the Aspects theory; but this is not so. Neither theory makes the predicate `takes up to two PPs' directly expressible - and we think this is as it should be, for the generalization does not deserve to be taken seriously: the real requirements of the verb `joke' are much more specific than that, and involve the type of PP (cf `joke with someone about something', *`joke of someone for something'). Hudson reads too much into the expository exemplification in our sketch.

On the other hand, both the Aspects theory and our own are capable of expressing Hudson's "linguistically simple fact"; and if one insists on trying to decide which expresses it better, the only thing we can think of that might be used to independently differentiate them, using principles that have been used elsewhere in linguistics, is the evaluation metric of orthodox generative phonology, which has simplicity measured by counting symbols. We do not subscribe to such a measure, but those who do may care

to note that the notation using the fewest independent symbols is ours. Thus Hudson has offered an entirely spurious argument for preferring the Aspects theory of subcategorization.

2. Constituent order

Hudson's comment on this topic is brief enough to quote in full.

> [Their] second point is that phrase-structure rules should not impose an order on the elements that they introduce, but should leave this job to a different set of rules, called 'linearization rules'. This idea has often been aired, and persuasive arguments have often been offered in favour of it, but it doesn't seem to have had a great deal of effect on most users of phrase-structure rules. I wish Gazdar and Pullum success in their campaign.

This comment leaves us wondering whether Hudson actually read our paper, or merely saw what topic we were about to address and guessed at the content. Consider this contrast:

> Hudson's paraphrase of Gazdar and Pullum:
>
> "phrase structure rules should not impose an order on the elements that they introduce"
>
> Gazdar & Pullum (p 108n):
>
> "syntactic representations always display both [dominance and precedence] simultaneously, and so do phrase structure rules"

Phrase structure rules originate in the theory of Post rewriting systems, in which they are defined in terms of strings of symbols. Our observation that they display precedence relations is thus merely analytic.

Furthermore, our paper makes no reference to anything called 'linearization rules', though Hudson actually puts this in quotes, as if it came from us. Various workers in differing versions of generative grammar have indeed proposed to have rules defining unlinearized structures which are mapped at a late stage of derivation into ordered structural descriptions of a more familiar sort. But the theory of constituent order developed in our paper has no such mapping, nor any level of unordered structural description, and thus we have no 'linearization rules' that "impose an order on the elements" in a constituent.

What we do in fact propose is a way of defining a grammar, i e, a set of phrase structure rules of the ordinary sort, in a way that allows generalizations about

linear precedence to be captured as general clauses in the definition. Grammars defined in this way are said to be given in 'ID/LP format'. But altering the definition of phrase structure rules is exactly what we do not do, and we have a long footnote (p 108) contrasting our proposal with numerous others on just this point.

3. The notion 'head'

The remaining comment concerns the notion 'head'. Here again we will quote Hudson in full:

> The third point is that it would be better to treat the notion 'head' as a basic category, expressed directly in the grammar, rather than as a derivative notion. Here too I hope their ideas will win an audience, though I should point out that this is a fundamental tenet of dependency theory, as Gazdar and Pullum know, but that they present it as though it was only recently introduced into theoretical linguistics.

Again we note that Hudson's account of our claims contains an exact contradiction of what we actually say:

> Hudson's paraphrase of Gazdar & Pullum:
>
> "The third point is that it would be better to treat the notion 'head' as a basic category"
>
> Gazdar and Pullum (p 113):
>
> "Note that the symbol H is not itself the name of a syntactic category ..."

Hudson is right in thinking that we know about the dependency grammar use of a concept to which the term 'head' (among various other terms) is sometimes attached. But he appears not to appreciate that the term 'head' (like, for example, the term 'marked') is one that linguists of many different persuasions employ, often in widely different senses. Zwicky (1985) distinguishes five distinct definitions of 'head' employed in morphology and syntax, and distinguishes the dependency grammar notion from all of them. X-bar syntax employs a notion of 'head' that sometimes coincides with the referent of the dependency grammar use of the term (e g in the case of VSO sentences) and sometimes does not (e g in the case of coordinate constructions). But such distinctions appear to be lost on Hudson. Once the word 'head' appears, it is a knee-jerk reaction for him to think 'dependency grammar'. But we do not employ an analog of the dependency-grammar notion of head at all.

Formally, there are major differences between our relation 'is the head of' and the one used in dependency

grammar. In our X-bar phrase structure theory, the relation 'is the head of' holds between daughter and mother nodes in a tree (sometimes the same term is also used to denote the ancestral of the head-of relation). In dependency grammar, relations of dependency hold only between words, and the notion of constituency crucial to every claim of our theory is not available. Our claims in the paper are not only independent of dependency theory, but incommensurable with it in a fairly profound sense.

The claims we make about the notion 'head', like the other claims in our paper, are entirely concerned with the statement of a phrase structure grammar. We do not employ any notion of 'dependency', and we do not posit 'head' as a category. What we do is to introduce a symbol H, interpreted as a copying variable in the sense of Hellan (1977: 90), into certain statements. These statements in part determine a phrase structure grammar via certain conventions, notably the head feature convention, which determines a category corresponding to the H in the rule by reference to feature specifications on the mother (left hand side) category, whose 'head' (in one sense) the H in effect represents. As before, we are left wondering whether Hudson actually read any of what we said about this idea and its applications.

4. Conclusion

Hudson begins his discussion of our paper by saying: "this paper makes three rather separate points, and I don't quite see why they were put together into a single article". Let us briefly explain. As Heny (1979) and Hendrick (1980) have pointed out, a major problem with the 'standard theory' view of subcategorization concerns lost generalizations about the linear order of sister constituents. Our ID/LP format for grammar definition provides an immediate solution to the problem of expressing the relevant generalizations about the order of the complements that a lexical item subcategorizes for, given the approach to subcategorization that we first outline in our paper. A remaining problem concerns the expression of the generalizations that can be made when the lexical head of a phrase precedes (as in English) or follows all of its subcategorized complements. The theory involving the H variable that we characterize in our paper provides a solution to that problem, inter alia.[2]

In our 1981 paper, (i) we replace Aspects-style context-sensitive subcategorization mechanisms by pointers that associate lexical (sub)categories with the rules that introduce them; (ii) we sever dominance from precedence in the rule definition system without changing the nature of the rules or the structures postulated; and (iii) we capture the notion 'head' without assuming any primitive category or relation called 'head'. Unlike Hudson, we see

these matters as closely and interestingly related.

Concluding his discussion of the book in which our paper appeared, Hudson wonders (p 450): "why is it so difficult (as I have shown) to grasp the main points of some of the papers?" We will not speculate here as to the answer to this question, but we hope that this note will assist Hudson, and perhaps others, in grasping the main points of at least one of the papers.

FOOTNOTES

* This note was prepared while Gazdar was a Fellow at the Center for Advanced Study in the Behavioral Sciences; he is grateful for the financial support provided by the Alfred P Sloan Foundation, the System Development Foundation, and the Economic and Social Research Council (U K). Pullum's research has been partially supported by a Faculty Research grant and by the Syntax Research Center at the University of California, Santa Cruz. Thanks to Bill Ladusaw for useful comments and suggestions.

1 This approach is taken in the pre-ID/LP work summarized in Gazdar (1982; written in 1980). It is not taken in Gazdar, Klein, Pullum and Sag (1985), as it happens.

2 There is a technical difficulty with that solution, pointed out by Shieber (1984), which has caused us to depart from its specific proposal; in Gazdar, Klein, Pullum and Sag (1985: 50), we offer a revised solution that has similar coverage but does not mention H in any LP rule (though H continues to appear in ID rules, of course, for reasons relating to feature distribution).

REFERENCES

Chomsky, Noam (1965) Aspects of the Theory of Syntax. Cambridge, Mass: MIT Press.

Gazdar, Gerald, and Geoffrey K Pullum (1981) 'Subcategorization, constituent order, and the notion "head"'. In Michael Moortgat, Harry v d Hulst, and Teun Hoekstra, eds The Scope of Lexical Rules. Dordrecht: Foris, 107-123.

Gazdar, Gerald (1982) 'Phrase structure grammar'. In Pauline Jacobson and Geoffrey K Pullum eds, The nature of Syntactic Representation. Dordrecht: Reidel, 131-186

Gazdar, Gerald, Ewan H Klein, Geoffrey K Pullum and Ivan A Sag (1985) Generalized Phrase Structure Grammar. Oxford: Blackwell; Cambridge, Mass: Harvard University Press.

Hellan, Lars (1977) "X'-syntax, categorial syntax and logical form". In Thorstein Fretheim and Lars Hellan, eds Papers from the Trondheim Syntax Symposium, Trondheim, 83-135.

Hendrick, Randall (1980) 'Subcategorization: its form and functioning'. Linguistics 19, 871-910.

Heny, Frank (1979) Review of Noam Chomsky, The Logical Structure of Linguistic Theory, New York: Plenum. Synthese 40, 317-352.

Hudson, Richard (1983) Review of Michael Moortgat, Harry v d Hulst, and Teun Hoekstra, eds The Scope of Lexical Rules, Dordrecht: Foris. Studies in Language 7, 439-50.

Sheiber, Stuart M (1984) 'Direct parsing of ID/LP grammars'. Linguistics and Philosophy 7, 135-154.

Zwicky, Arnold M (1985) 'Heads'. Journal of Linguistics 21, 1-29.